

Pembangkitan Data dari Distribusi Log-logistik

ACENG KOMARUDIN MUTAQIN, ABDUL KUDUS

Program Studi Statistika, Universitas Islam Bandung
Email: aceng.k.mutaqin@unisba.ac.id

ABSTRAK

Distribusi log-logistik merupakan salah satu distribusi yang dapat diaplikasikan untuk data mutu lingkungan. Makalah ini membahas pembangkitan data dari distribusi log-logistik mulai dari pembangunan algoritme sampai pembuatan program komputer pada perangkat lunak MATLAB. Metode pembangkitan datanya menggunakan metode inverse transform. Ada dua jenis data yang dibangkitkan, yaitu data lengkap dan data tersensor kiri. Dalam makalah ini diberikan contoh data yang dibangkitkan dari distribusi log-logistik.

Kata Kunci: distribusi log-logistik, metode inverse transform, data lengkap, data tersensor.

1 Pendahuluan

Distribusi log-logistik merupakan salah satu distribusi yang dapat diaplikasikan untuk data mutu lingkungan (Warsono, 1996), data asuransi (Klugman dkk, 2004), data survival (Collet, 2003), dan data ekonomi (Risk, 1961). Ada beberapa bentuk fungsi densitas dari distribusi log-logistik, hal ini tergantung pada jenis aplikasinya.

Beberapa perangkat lunak, ada yang belum memuat perintah untuk membangkitkan data dari distribusi log-logistik. Salah satu diantaranya adalah perangkat lunak MATLAB. Tujuan dari makalah ini adalah membahas pembangkitan data dari distribusi log-logistik mulai dari pembangunan algoritme sampai pembuatan program komputer pada perangkat lunak MATLAB. Metode pembangkitan data yang digunakannya adalah metode *inverse transform* (Law & Kelton, 2000). Ada dua jenis data yang dibangkitkan yaitu data lengkap, dan data tersensor kiri.

2 Distribusi Log-logistik

Salah satu distribusi yang dapat diaplikasikan untuk data mutu lingkungan adalah distribusi log-logistik diperumum (Warsono, 1996). Distribusi log-logistik adalah distribusi khusus dari distribusi log-logistik diperumum, dimana bentuk fungsi densitasnya adalah

$$g(x; \alpha, \beta) = \frac{\alpha}{x} \left[\frac{e^{\beta x^{\alpha}}}{(1 + e^{\beta x^{\alpha}})^2} \right]; x > 0,$$

dimana $\alpha > 0$ adalah parameter skala, dan $-\infty < \beta < \infty$ adalah parameter lokasi. Momen ke- k untuk distribusi log-logistik (Klugman dkk., 2004) di atas adalah

$$E[X^k] = e^{-k\beta/\alpha} \Gamma\left(1 + \frac{k}{\alpha}\right) \Gamma\left(1 - \frac{k}{\alpha}\right); -\alpha < k < \alpha,$$

Sedangkan fungsi distribusinya adalah

$$G(x; \alpha, \beta) = \left(\frac{e^{\beta x^{\alpha}}}{1 + e^{\beta x^{\alpha}}} \right); x > 0.$$

3 Algoritme Pembangkitan Data

Pada bagian ini akan dibangun algoritme untuk membangkitkan data lengkap dan data tersensor kiri dari distribusi log-logistik. Misalkan $u = G(x; \alpha, \beta)$, atau

$$u = \left(\frac{e^{\beta x^{\alpha}}}{1 + e^{\beta x^{\alpha}}} \right),$$

sehingga

$$x = \left[\frac{u}{(1-u)e^{\beta}} \right]^{1/\alpha}. \quad (1)$$

Berdasarkan uraian di atas, dengan menggunakan metode *inverse transform* algoritme untuk membangkitkan data lengkap dari distribusi log-logistik adalah sebagai berikut:

Tentukan nilai parameter α dan β .

- (1) Bangkitkan data u dari distribusi seragam (0,1).
- (2) Hitung nilai x dengan menggunakan Persamaan (1).

Sementara itu, dengan mengacu pada Warsono (1996), algoritme untuk membangkitkan data tersensor kiri dari distribusi log-logistik adalah sebagai berikut:

- (1) Tentukan nilai parameter α dan β .
- (2) Tentukan persentase data tersensor (p).
- (3) Bangkitkan data dari distribusi log-logistik dengan menggunakan algoritme untuk data lengkap.
- (4) Urutkan data dari kecil ke besar.
- (5) Tentukan titik sensor kiri sedemikian sehingga persentase data yang kurang dari titik sensor tersebut sebesar p .

4 Program MATLAB dan Contoh Data

Program MATLAB untuk membangkitkan data lengkap dari distribusi log-logistik disajikan dalam Gambar 1. Sedangkan Gambar 2 mengilustrasikan program MATLAB untuk membangkitkan data tersensor kiri dari distribusi log-logistik.

Dengan menetapkan nilai $\alpha = 4,7110$ dan $\beta = -6,3344$, hasil pembangkitan data lengkap sebanyak $n = 30$ pengamatan disajikan dalam Tabel 1. Sementara itu dengan menetapkan $\alpha = 4,7110$, $\beta = -6,3344$, $n = 90$, persentase sensor untuk 6 kelompok data $p = 0,3; 0,25; 0,2; 0,15; 0,1; 0,05$, data tersensor kiri hasil pembangkitan dengan menggunakan program MATLAB yang ada pada Gambar 2 disajikan dalam Tabel 2. Titik sensor untuk setiap kelompok masing-masing adalah 3,75; 3,16; 3,02; 2,75; 2,02; dan 1,86.

```
function X = GenerateLL(n,a,b)
% a = alpha dalam model TLL
% b = beta dalam model TLL
% program membangkitkan data berukuran n
% berdistribusi Log-Logistic, LL(a,b)
% rand('state',100)
for i=1:n
    U=rand(1,1);
    X(i,1)=(U/((1-U)*exp(b)))^(1/a);
end
```

Gambar 1. Program MATLAB untuk Membangkitkan Data Lengkap dari Distribusi Log-logistik

```
function [X,BD,t]=SkemaA(N,a,b,p)
% Program untuk membangkitkan data skema A
% N adalah sample size
% a dan b adalah parameter distribusi LL
% p adalah vektor persentase data yang tersensor

% Pembangkitan Data Distribusi LL
A = GenerateLL(N,a,b);
B = rand(N,1);
C = [A,B];
D = sortrows(C,2);
baris=N/6;
MAT=reshape(D(:,1),baris,6);
datasort=sort(MAT);
sensor=floor(N/6*p)+1;
t=sensor-1;
% Pemilihan BD
for i=1:6
    BD(i)=datasort(sensor(i),i);
end
X1=datasort(sensor(1):baris,1);
X2=datasort(sensor(2):baris,2);
X3=datasort(sensor(3):baris,3);
X4=datasort(sensor(4):baris,4);
X5=datasort(sensor(5):baris,5);
X6=datasort(sensor(6):baris,6);
X=[X1;X2;X3;X4;X5;X6];
```

Gambar 2. Program MATLAB untuk Membangkitkan Data Tersensor Kiri dari Distribusi Log-logistik

Tabel 1 Data Lengkap Hasil Pembangkitan

5,25	2,69	4,40
6,20	8,06	1,91
2,55	7,42	5,54
6,33	3,79	6,73
4,30	5,15	4,50
2,39	2,62	4,89
3,13	3,59	4,81
3,99	6,37	3,50
7,43	5,10	4,40
7,75	7,51	2,75

Tabel 2 Data Tersensor Kiri Hasil Pembangkitan

<3,75	3,33	<3,02	4,85	4,36	4,13
7,43	<3,16	5,25	3,64	2,62	4,89
4,00	3,36	7,18	2,75	4,50	3,13
6,39	3,80	3,59	4,40	3,46	5,32
<3,75	4,87	5,12	4,07	4,88	3,66
6,20	<3,16	7,42	3,99	6,37	2,60
3,75	4,50	7,75	7,50	5,10	3,86
5,25	<3,16	3,99	5,99	4,86	3,46
3,83	5,15	4,40	<2,75	7,51	6,73
4,40	3,16	5,38	3,79	2,02	4,30
<3,75	4,93	<3,02	<2,75	6,33	2,85
4,57	4,13	3,05	4,81	5,46	4,59
6,63	5,54	3,26	4,64	8,06	3,04
4,62	3,37	<3,02	3,13	<2,02	1,86
<3,75	7,52	3,02	4,24	2,39	3,50

Daftar Pustaka

- [1]. Klugman, S.A., Panjer, H.H., dan Willmot, G.E. (2004): *Loss Models: From Data to Decision*. New York, John Wiley & Sons.
- [2]. Law, A.M., dan Kelton, W.D. (2000): *Simulation Modeling and Analysis*, edisi ketiga, New York, McGraw-Hill.
- [3]. Warsono, (1996): *Analysis of Environmental Pollutant Data Using Generalized Log-logistic Distribution*, *Dissertation at University of Alabama at Birmingham*.
- [4]. Collet, D. (2003): *Modelling Survival Data in Medical Research*. London, Chapman and Hall.
- [5]. Risk, P.R. (1961): *The Graduation of Income Distributions*. *Econometrica*, 29(2), 171-185.